Haile Quals Day 2

Many of your papers use computational models to understand and interpret learning in humans. Please describe (a) How have these models been used to study learning; (b) What are their main advantages over purely behavioral research? (c) what are some of their disadvantages? Then describe the factors one might need to consider when creating or evaluating a model of complex skill learning.

Computational models are powerful tools that allow us to concretize and flexibly test theories of cognition that are more abstracted or are difficult to measure directly. Computational models of learning have been used to explain learning behavior, predict learning outcomes, and estimate reliable parameters for latent variables that might explain learning, individual differences and brain activity related to these specific cognitive behaviors.

Learning and memory in the brain result from an amalgam of differential computations that occur in highly interactive networks. Several theories from both human and animal experiments have allowed us to build both simple characterizations like reinforcement learning and large computational architectures like ACT-R which flexibly instantiate and represent those learning and memory theories for application towards novel tasks, explain aspects of learning we cannot measure directly, generate new theories of how learning progresses and even relate, more reliably, brain activity to behavior.

Using a set of relatively simple models, Collins (2018) demonstrated how different cognitive systems might interact during learning in a changing environment. The study also clearly demonstrated an advantage of cognitive modelling where different theories or hypotheses could be compared by using only one carefully designed task and one dataset.

Different environmental conditions or demands require that an agent learn and respond quickly, or slowly acquire a skill that needs to be robust over the long term. Behavioral and neuroimaging studies have documented that fast declarative and slower non-declarative learning systems are deployed in parallel and compete for control of behavior (Poldrack et al. 2003; McDonald and Hong, 2013). More traditionally, the more flexible and fast declarative system competes for and gains control of behavior first and inhibits the non-declarative system preventing it from directing behavior (McDonald and Hong, 2013). Collins (2018) forwarded models that demonstrated how the interactions between these systems could change when the environment poses different demands.

Collins (2018) sought to test how working memory (arguably learning through PFC + MTL based declarative memory, rather than solely through working memory) and Reinforcement learning interacted and controlled behavior in an environment with different demands on capacity and durability of learning. The study used a simple feedback-based, deterministic, stimulus-response learning task with two critical design elements: blocks that vary in the number of objects that must be learned (3 items to associate with 3 responses versus 6 items to associate with 3 responses), and a surprise test after a long distracting break, without feedback. This experiment was designed to tease apart key aspects of these learning systems: learning speed, capacity limitations and durability of learned associations. Working memory is known for fast learning but is capacity limited and learned associations are not robust to interference and long delays. In comparison, reinforcement learning acquires associations slowly but durably, and is virtually unlimited in capacity.

Several computational modelling advantages were leveraged by this study. For instance, some cognitive dynamics are too intricate and occur too quickly to be reliably captured by neuroimaging techniques to be correlated to behavior in humans. Robust models like Reinforcement Learning, that have been extensively studied through more invasive physiological experiments, in animal models could be applied to behavioral data and prove (or disprove) if a specific cognitive computation had guided behavior in response to specific environmental learning demands.

Collins (2018) clearly defined its learning models using a series of single-component reinforcement learning (RL) and working memory (WM) only models, and 2-component, WM and RL, models. The 2-component models are broken up into independent and interacting models. The RL models are based on classical Q-learning (where actions for specific states are assigned a value, Q, based on the amount of expected reward they are likely to receive) RL models which uses a softMax choice policy and the all-important reward prediction error (d) that signals the value of an action by computing the difference between given and expected reward. A set of modifications were applied by the researchers to further capture features of human performance like forgetting, perseverance and initial bias. This resulted in a complex model with multiple parameters (learning rate, softMax temperature only for the simplest model, with additional parameters like decay-rate and perseverance to represent the additional cognitive and behavioral features). The WM model, similarly, utilizes a Q-learning model (now called W-learning) with critical changes that captures WM features. For instance, the learning rate parameter, which weights how much influence previous rewards have on the current state-action pair, was set to 1 to represent perfect recall of the reward. An additional decay parameter also distinguishes the WM model from RL by heavily discounting Q-values to simulate the rapid decay of working memory contents. According to their theoretical premise, the mixture models are also sensitive to set size and compete at the level of choice policy only in the independent models, and the WM portion contributes to RL reward prediction errors in the interacting version. Then, by fitting their set of single and dual-system models to behavioral data, Collins and colleagues showed that an interacting WM and RL characterization of learning most likely resembled how subjects learned the associations as opposed to utilizing RL only or WM only single system for learning, which is largely consistent with others' cognitive models of learning (e.g., Poldrack et al. 2003).

The design of the models summarized above exemplifies the flexibility, concreteness and breadth of behavioral and cognitive features that could be modelled and explained. In a series of papers Collins and colleagues have demonstrated how WM and RL might trade off during learning (Collins, 2018), decoded EEG markers of these learning mechanisms by correlating with parameter model estimates (Collins and Frank, 2018), and even characterized WM - RL interaction impairments in disorders like Schizophrenia (Collins et al., 2014) further demonstrating the fact that cognitive modelling techniques could be applied in clinical settings.

Collins and Colleagues have used a relatively simple yet powerful set of models to explain cognitive dynamics in a simple task. But the vast majority of human learning experience occurs in more complex environments and may perhaps need more comprehensive models to model, explain, and predict behavior. I am not arguing that complex environments always need complex models (there really isn't even an objective barometer for measuring complexity) but just that, sometimes, larger models that capture more pieces of human cognition like perception, motor planning, executive function, conflict resolution and motivation in a theoretically unified manner may offer more insights. I am obviously trying to illustrate the merits of taking advantage of an integrated, modular architecture like ACT-R (Adaptive Control of Thought-Rational; Anderson, 1983; Anderson and Lebiere, 1998).

The ACT-R cognitive architecture integrates several theories of sensation and perception, motor planning and execution and attentional activation with two critical learning and memory systems, a declarative store of memories and a procedural module that learns and performs skills primarily through reinforcement learning, to model learning and responding behavior. The ACT-R architecture is almost infinitely flexible and has been used to model a wide variety of both complex (Lebiere and Anderson, 2001) and simple tasks (Lovett et al. 2000), estimate reliable parameters of very specific cognitive and biological properties (Xu and Stocco, 2021), even evaluate popular learning strategies (Marewski and Schooler, 2013), to mention a few.

Behavioral experiments are at the center of psychological studies and have been for a long time. But behavioral experiments are sometimes criticized for lacking reliability, poorly explaining, or revealing latent variables, and failing to characterize individual differences (Rouder and Haaf, 2019; Xu and Stocco 2021). The addition of computational models to behavioral experiments relieves some of these shortcomings and increase reliability and reveals replicable latent variables.

For instance, to address the lack of reliability in subject results in the Probabilistic Stimulus Selection Task (PSS) that arises due to its design, Xu and Stocco (2021) built a cognitive model of the task based on ACT-R and estimated stable parameters that represented the D1 and D2 basal ganglia pathways. The PSS task trains subjects, with feedback, which Japanese characters have high chance of being correct and which characters are more likely to be incorrect. Two of the tasks dependent variables, "choose" and "avoid" accuracies signal an individual's reliance on the D1 dopamine receptor mediated direct pathway and the D2 dopamine receptor mediated indirect pathway, respectively. Xu and Stocco, (2021) showed that estimations of the choose and avoid accuracies were poorly replicated in a test-retest experiment. However, parameters for D1 and D2, computed by their ACT-R model, which assumed that that the PSS task is learned solely through reinforcement learning, estimated for each participant in both sessions were significantly correlated. In other words, they have managed to recover reliable parameters from noisy behavioral data.

There are large implications here for individual differences studies too. The study has managed to extract reliable, individualized characterizations of cognitive and biological variables that could be used to explain or predict other related behavior. Currently most behavioral experiments are poor at capturing individual differences primarily because most experiments are optimized to differentiate group differences (Rouder and Haaf, 2019), much like the PSS task, was designed to do (Xu and Stocco, 2021).

Other ACT-R based studies have successfully estimated individual parameters of cognition that were then used to reliably predict individual performance in a different task. Lovett, Daily and Reder (2000) estimated a parameter for working memory capacity in ACT-R. Individual difference in working memory capacity are highly documented and predictive of other cognitive features like general intelligence. Lovett et al. (2000) used two working memory tasks, a digit span task and the n-back task. They demonstrated that a single parameter, that is a working-memory analog in ACT-R (represented by the source activation (W), which is part of the equation that determines whether a chunk of information will be activated or not in LTM), estimated from a model of the digit span task reliability predicted individual performance in the n-back task. This is very interesting as behavioral results from span tasks and n-back tasks are normally not correlated even while they purport to measure aspects of working memory (Redick and Lindsey, 2013).

Computational models are powerful flexible tools and have pushed the fields of psychology and neuroscience by leaps and bounds, (Wilson and Collins, 2019), however, like most other tools and scaffolds for thinking, they have their disadvantages. The most obvious of these is perhaps the fact that we can only model and test a select few things *we know to test*. There will, probably, always be a model out there that would explain our phenomena of interest better. To dispense with other obvious disadvantages, computational models require acquiring more specialized mathematical and computational knowledge. They might also be resource intensive in-terms of computations. For instance, simulating many subjects performing hundreds of trials, while producing powerfully reliable data, might take hours or weeks (or months) to complete, depending on available processing power higher levels of which might be expensive to obtain.

Perhaps more serious disadvantages with computational models that might arise are difficulty in interpretation of findings as models might be poorly designed, too simplistic, or too complex. Additionally, sometimes, a model could explain behavior well but, so could other models that were built under different theoretical assumptions. For instance, Haile, Prat and Stocco (2020) have demonstrated that declarative long-term memory, which shares a feature of high durability with reinforcement learning might explain some or a lot of the learning in Collins' stimulus response task, that Collins (2018) claims could only be due to reinforcement learning. I understand that these are criticisms that may apply to a broad set of tools for scientific inquiry, but they are relevant here as there is significant difficulty in designing and building computational models.

Computational models are precise mathematical characterizations of hypotheses to be tested. I suppose the first and most important directive for building a model is to build a *good* model. A good model captures the complexity (or simplicity) of the system being modeled, is easy to interpret, and makes sound assumptions about the system that could be clearly demonstrated and tested. Additionally, a good model is only as good as the experimental design (Wilson and Collins, 2019). Capturing and modelling poor behavioral results from poorly designed experiments will not tell us anything more than how bad the design was. Therefore, a first step in evaluating how good the model is to inspect how good the experimental design was that produced the behavioral data.

Attempting to model complex skill learning is a daunting task. We would describe some skills as complex because they have several environmental parameters that vary widely, have complex multi-level, hierarchical or multi-component, broad concepts. We can argue that even the simplest of tasks require multiple learning and memory systems to learn but this is especially true for complex tasks. There will often be multiple ways of learning complex skills which would involve differential combinations of more fundamental learning mechanisms. When confronted with such complexity, the task of modelling and explaining intricate cognitive dynamics or individual differences will seem very daunting. Therefore, it is important to think very carefully about the aspects of the complex skill one wishes to capture and think carefully about the parts that might overlap with simpler skills (if so, other models could be conglomerated to explain an hierarchically more complex aspect of the skill. For example, if we were to model learning the python programming language, it might be imprudent to construct models that mimic learning starting from individual keypresses (unless that was the goal of course). The point is, thinking clearly about the goal will provide much needed constraints that make the task tractable but will also have other benefits. Clearly laid out hypotheses will make interpretation of the results from the models easier and, perhaps (hopefully) impactful. For example, is the goal to find individual differences? Is the goal to disprove theories for learning that apply more broadly? or both? Lastly, careful consideration of applicable theories for skill learning would naturally provide useful perspectives and scaffolds. For instance, the model for skill development described by Anderson and others, that shows the 3-stage transformation of skills from 'encoding phase' to the more automatic 'responding phase' Tenison et al., 2016; Anderson, 1982) are well defined and testable but there may be other less well-known models of skill acquisition.